

Supplemental Materials:

- 1) **Materials and Methods**
- 2) **Supplemental Figures**
- 3) **Supplemental Table Legends**
- 4) **Author Contributions**
- 5) **Supplemental Acknowledgements**

Materials and Methods:

Primary HSPC Isolation, Viral Transduction, *In Vitro* self-renewal, Flow Cytometry, ChIP, and RNA analysis.

Mice were euthanized in accordance with Medical College of Wisconsin Institutional Animal Care and Use Committee guidelines (IACUC# AUA2688). HSPCs were isolated from both male and female animals during the study. No animals or biological replicates were excluded from the analyses. HSPC isolations and viral transductions were performed as previously described{Milanovich:2015dt}. Lentiviral shRNAs used in this study came from the TRC consortium: *Rad21*-shRNA (shRNA1: TRCN0000174832; shRNA2: TRCN000017684), *Hoxa9*-shRNA (shRNA1: TRCN0000012510; shRNA2: TRCN0000012508) *Hoxa7*-shRNA (shRNA1: TRCN0000225923; shRNA2: TRCN0000218002), *Ctcf*-shRNA (shRNA1: TRCN0000039019; shRNA2: TRCN20000039022) or empty vector constructs (pLKO.1). To maintain selection, 1 ug/mL of puromycin and 100 ug/mL of hygromycin was used in methylcellulose culture. Methocult, RNA isolation, real-time PCR, Chromatin immunoprecipitation (ChIP), and Flow Cytometry was performed as in{Milanovich:2015dt, Pulakanti:2013jx, Stelloh:2016dy}. Two independent investigators performed assessment of colony forming units. Primers used in gene expression studies are shown in Table S1.

Immunoprecipitation:

32-D murine immortalized myeloblast cells were collected and washed with ice cold PBS. Cells were suspended in IP lysis buffer (Thermofisher Scientific Cat# 87787) containing PMSF and protease inhibitor cocktail. Whole cell lysates were isolated and 2000 µg of protein was used for immunoprecipitations. Immunoprecipitations were performed as previously described {Zhang:2008be}. Briefly, samples were suspended in 1200 µL of IP lysis buffer and precleared with 50 µL of Protein G PLUS-Agarose beads (Santa Cruz Biotechnology, Cat# sc-2002) and 1 µg normal mouse IgG antibody (Santa Cruz Biotechnology, Cat# sc-2025) for 1 hour at 4°C. 10 µg of antibodies or IgG were then added to the lysates and they were incubated at 4°C for 2 hours. 50 µL of beads were then added to the lysates and rotated overnight at 4°C. The following morning, agarose beads were washed 5 times with 1 mL IP lysis buffer. Antigens were eluted from the beads using SDS-Sample buffer (Boston BioProducts, Cat# BP-111R). Samples were boiled for 10 minutes at 95°C and spun for 3 minutes at 2500 RPM. Equal volumes of samples were run on 8% SDS-PAGE gels along with an input control, and the presence of RAD21 was evaluated using anti-RAD21 antibody (Abcam, Cat# ab992). 293T co-immunoprecipitations were performed as previously reported {Rao:2010do}. All immunoblotting was performed as previously {Milanovich:2015dt, Pulakanti:2013jx}. Antibodies used for protein applications are summarized in Table S2.

DNA Content Determination:

Cells were collected, washed with bone marrow wash medium, and counted using a hemocytometer. Per sample, 500,000 cells were suspended in PBS and stained with propidium iodide (PI). Percentages of cells in each stage of the cell cycle were determined by flow cytometry using an Accuri C6 flow cytometer. Analysis was performed using FCS Express software. Statistical analyses were determined by Student's t-test (two tailed, unpaired).

Microarray:

RNA from cells at the end of primary plating were collected and isolated as above. Arrays were run for 2 biological replicates each of empty vector control, *Rad21*-shRNA1, and *Rad21*-shRNA2. Probes were generated and hybridized to Affymetrix Mouse 430A_2 chips. CEL files were normalized with RMA {Irizarry:2003ge} and subsequently processed using Genesoft {Reich:2006jy} and GSEA {Subramanian:2005jt} similar to our previous work using default parameters {Rao:2010do}. For identification of genes differentially expressed by *Hoxa9* overexpression, we downloaded and reanalyzed data from {Muntean:2010ir} (GSE 212299). Specific datasets used are listed in Table S3. All new microarray data have been deposited into the NCBI GEO omnibus (ID GSE76320).

Reanalysis of Cancer Genome Atlas datasets:

Data for all AML samples within the TCGA (<https://tcga-data.nci.nih.gov/tcga/dataAccessMatrix.htm>) were downloaded. Level 3 expression data and mutation identification were directly taken from TCGA. Statistical significance was determined using a Mann-Whitney U test.

RNA Sequencing Bioinformatics

RNA-seq data for cells infected with shRNAs to luciferase, *Rad21*, or *Smc3* (Mullenders:2015fi) (GSE 72368) were downloaded and reanalyzed as previously described (Pulakanti:2013jx). Briefly, raw data were downloaded and aligned to the mouse genome (mm9) using TopHat, and FPKMs for well-annotated genes were assigned using CuffLinks (Trapnell:2012kp). FPKMs for each gene are listed in Table S4.

Statistical Analyses

Statistical significance was determined by Student's t-test (two tailed, unpaired) for all figures except Figure 3C which was a Mann-Whitney test to account for the unequal variance between groups. The number of biological replicates is listed in the figure legend for each figure. Error bars in all bar graphs represents the standard error of the mean of biological replicates. Asterisks (*) represent p-values < 0.05 and crosses (†) indicate p-values < 0.01.

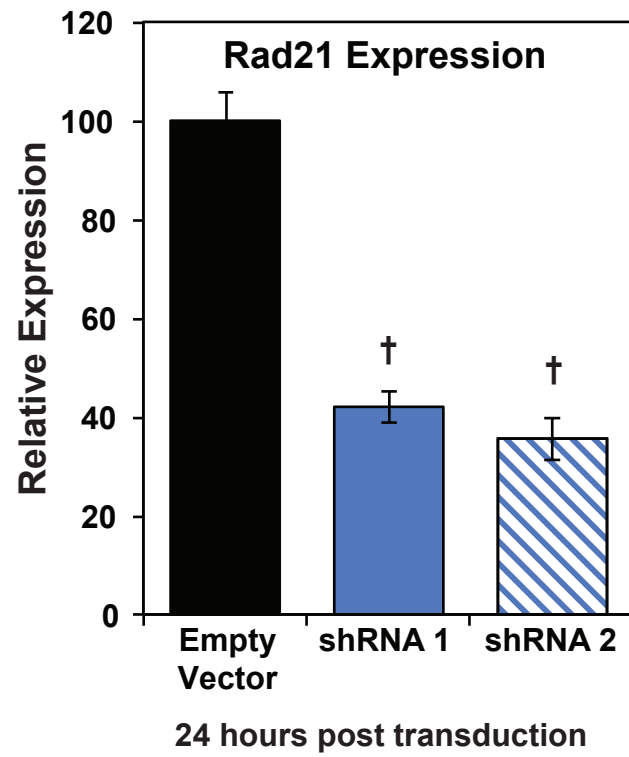


Figure S1. Lentiviral shRNA reduces *Rad21* levels in primary HSPCs. RT-qPCR analysis of primary HSPCs 24 hours after treatment with shRNA. Cells transduced with lentiviral shRNA were selected for 24 hours and RNA from puromycin-resistant cells was converted to cDNA, and analyzed using primers specific for *Rad21* mRNA. Each data point represents 7 independent biological replicates.

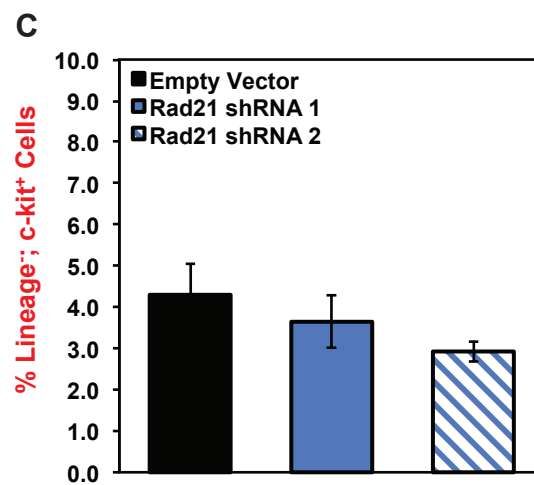
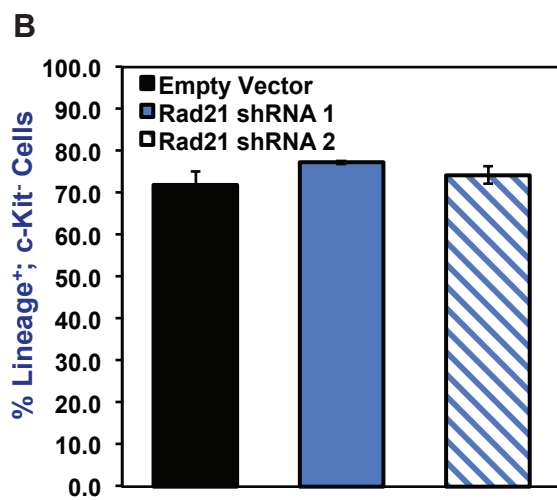
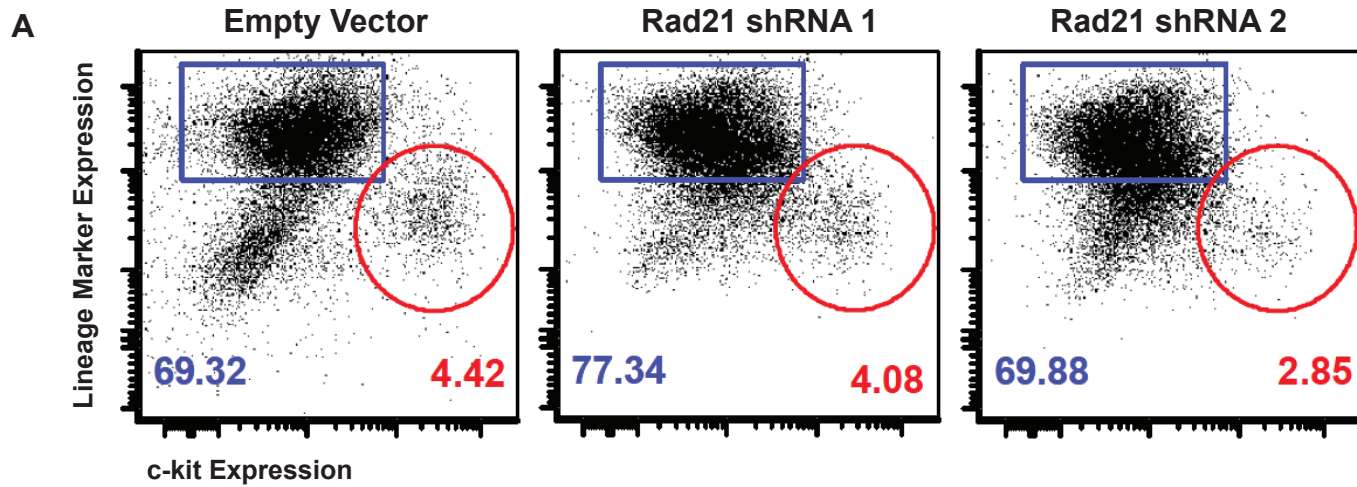


Figure S2. *Rad21* depletion does not selectively deplete HSPCs or differentiated lineages. (A) Representative flow cytometry dot plots of lineage marker (y-axis) vs c-kit (x-axis) performed on empty vector and *Rad21*-depleted HSPCs at the end of primary methylcellulose culture. (B) Bar graphs showing the percentages of c-kit-negative/lineage-positive averages from 3 independent biological replicates of the flow cytometry performed in A. (C) Bar graphs showing the percentages of c-kit-positive/lineage-negative averages from 3 independent biological replicates of the flow cytometry performed in A.

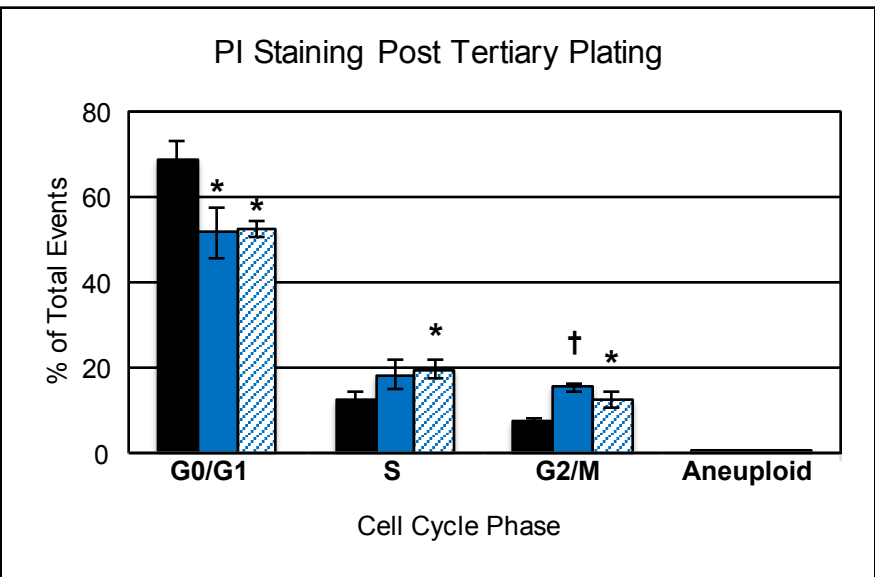
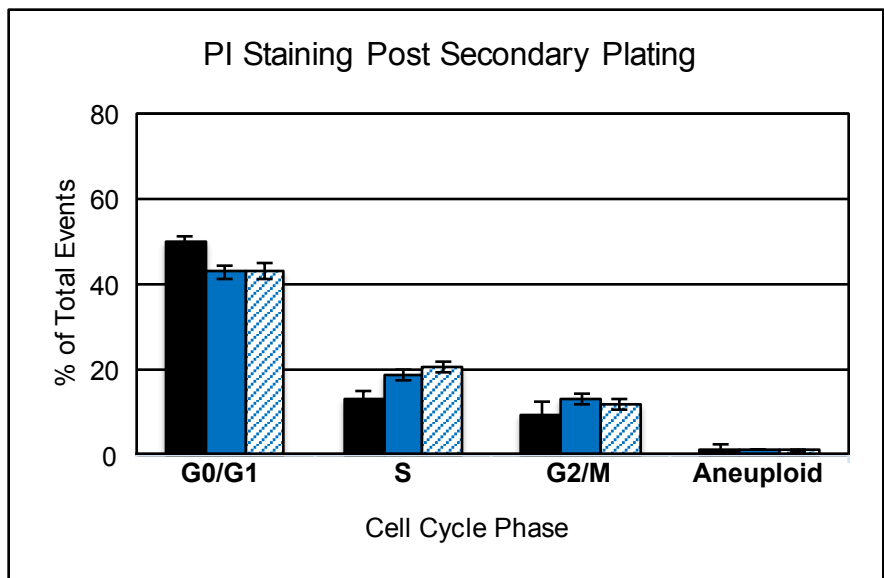
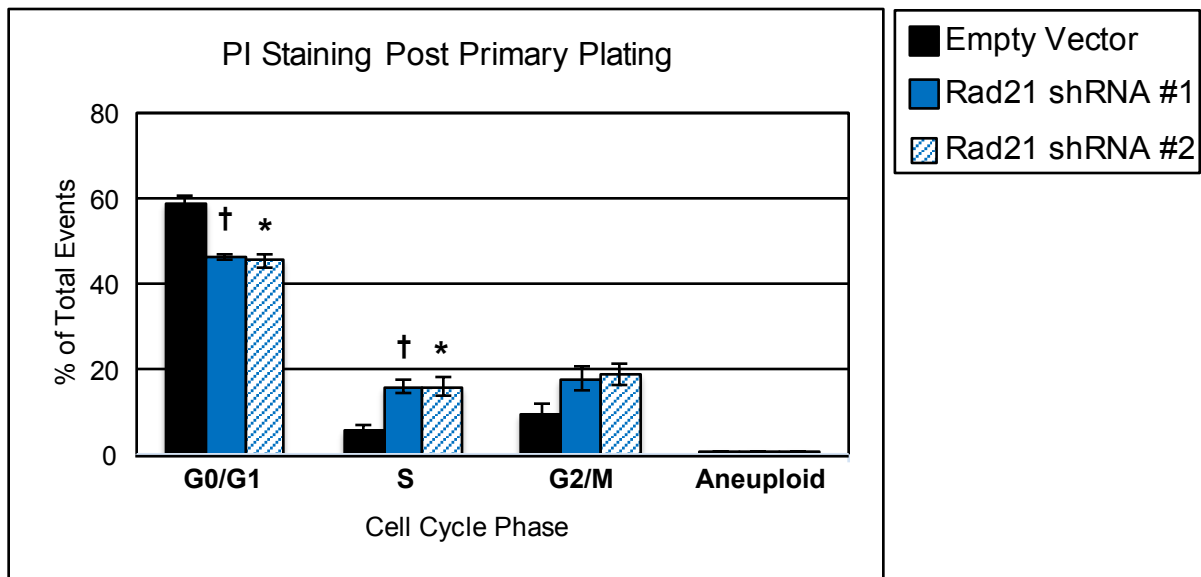


Figure S3. *Rad21* depletion alters cell cycle and does not induce aneuploidy. Flow cytometry of propidium iodide stained control and *Rad21*-shRNA1 and shRNA2 HSPCs at the end of primary, secondary, and tertiary methylcellulose passages. All data points represent 3 independent biological replicates.

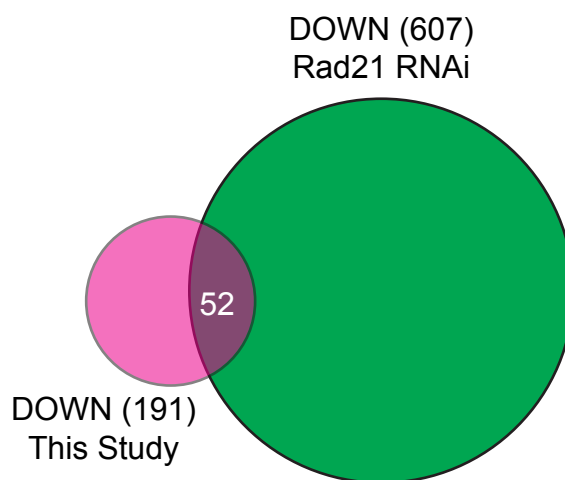
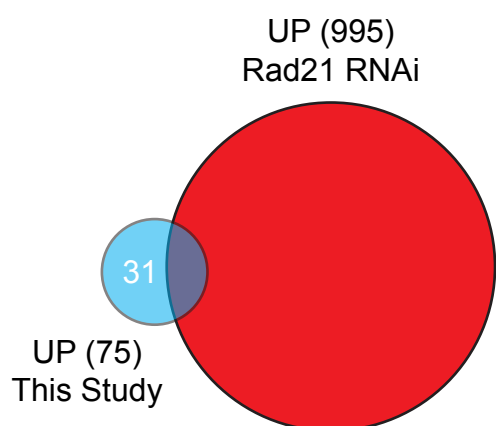
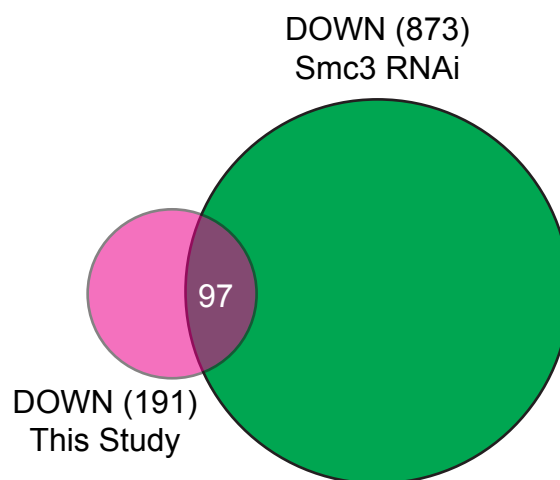
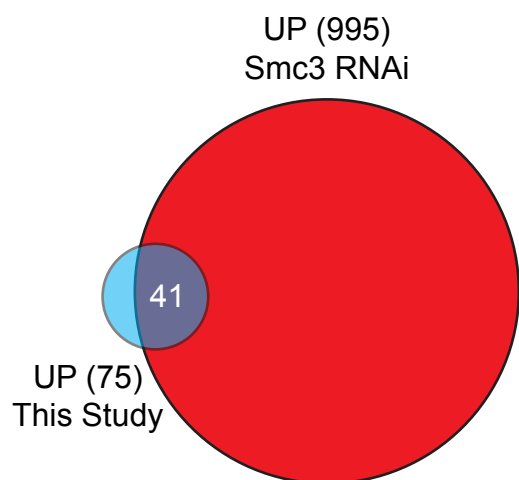
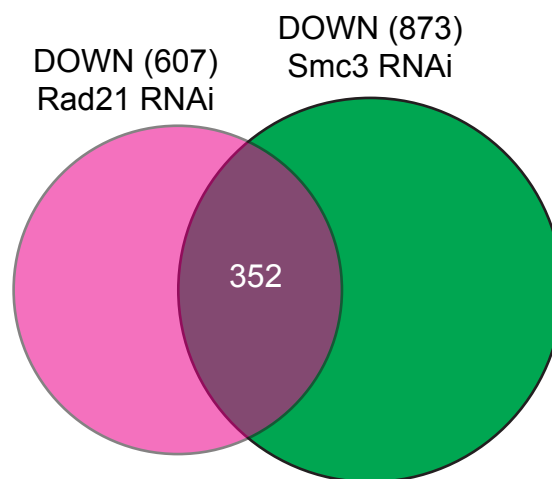
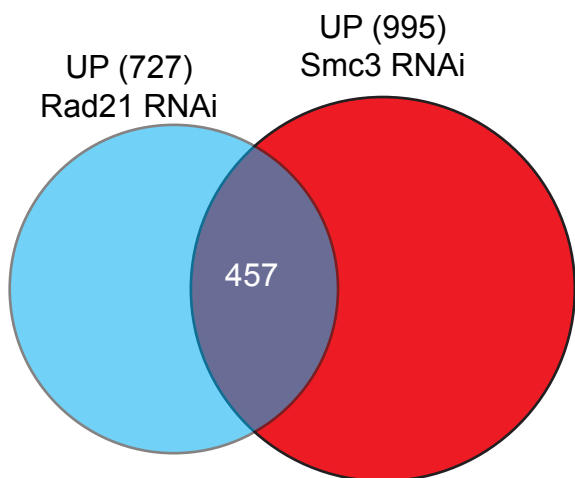


Figure S4. Overlap between transcriptome analyses of *Rad21*-depleted HSPCs vs previously published RNA-seq. Affected gene lists from *Rad21*-depleted HSPCs at the end of primary methylcellulose culture from our study and a previously published data set {Mullenders:2015fi}. Genes which showed significant changes (UP or DOWN) after *Rad21* depletion from this study (list of genes are shown in Table S5). For comparison, we used a similar fold-change for RNA-seq in which either *Rad21* or *Smc3* were depleted by RNAi (list of genes are shown in Table S4 {Mullenders:2015fi}). The number of genes in each list is shown in parentheses, and the size of each circle is proportional to the # of genes. Overlap between different gene sets is shown. Non-overlapping genes are likely secondary to subtle changes in the experimental conditions (differences in shRNA knockdown efficiency).

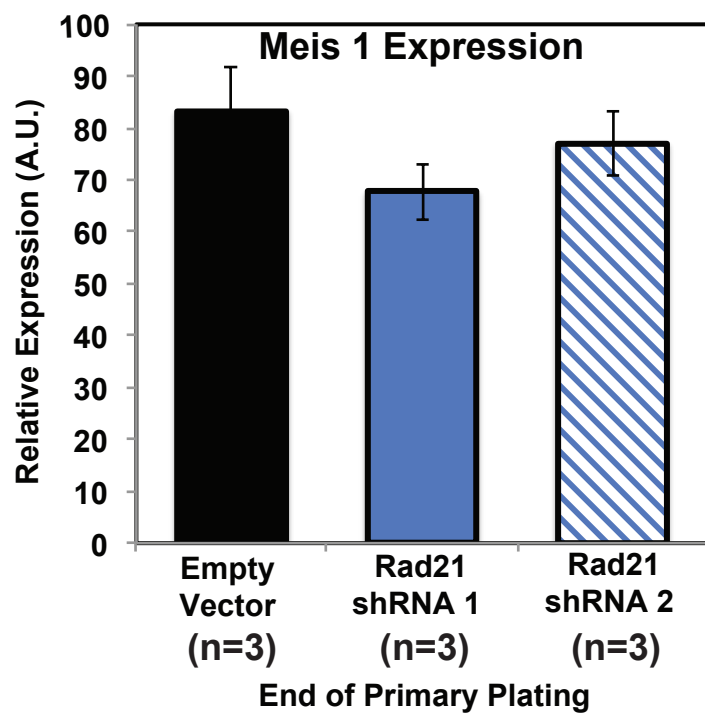


Figure S5. *Rad21* depletion does not affect *Meis1* expression. RT-qPCR results showing expression of *Meis1* in post primary methylcellulose culture *Rad21*-depleted or empty vector treated HSPCs.

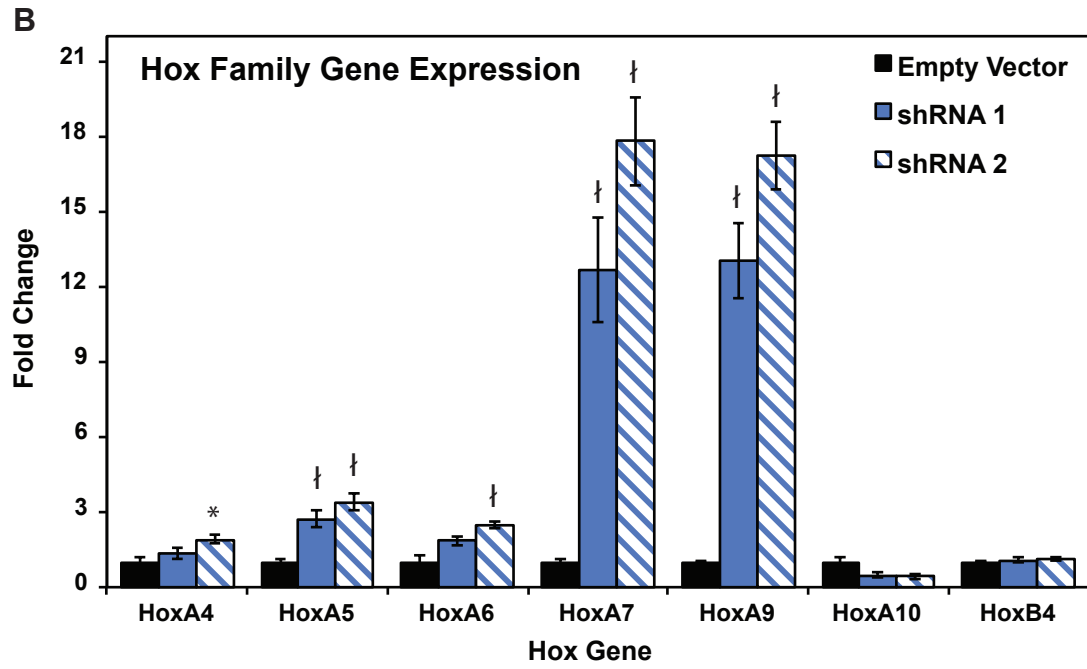
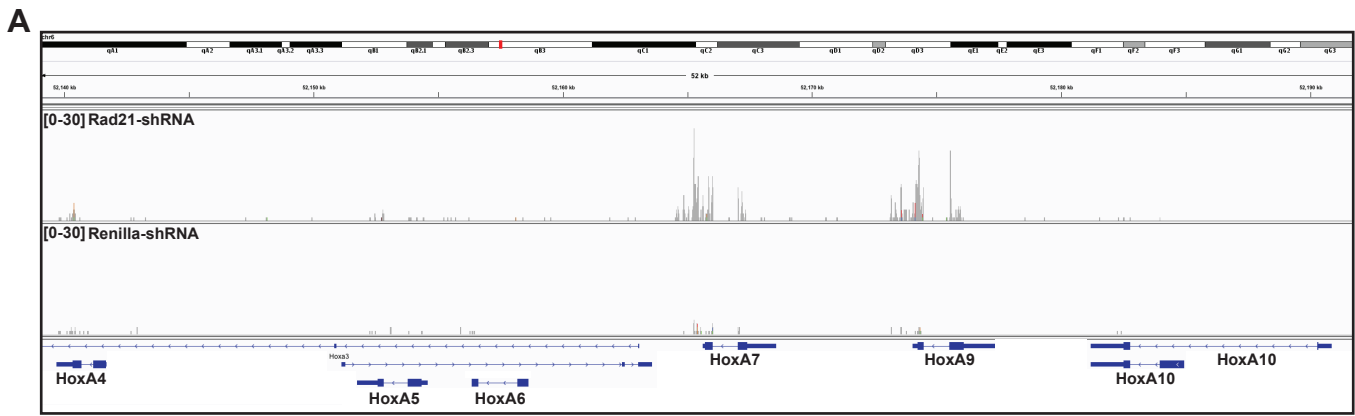


Figure S6. *Rad21* depletion increases expression of other genes at the *HoxA* cluster. (A) Map of the *Hoxa* genomic locus showing the 5' to 3' orientation of the *Hoxa10-Hoxa4* genes. RNA-seq data is shown in *Rad21*-shRNA or Renilla-shRNA infected HSPCs. (B) RT-qPCR of genes in the *Hoxa* cluster as well as the unrelated *Hoxb4* gene. Each data point represents 6 independent biological replicates. The *Hoxa9* expression data included in panel B is the identical data shown in Figure 2C and was included here for comparison.

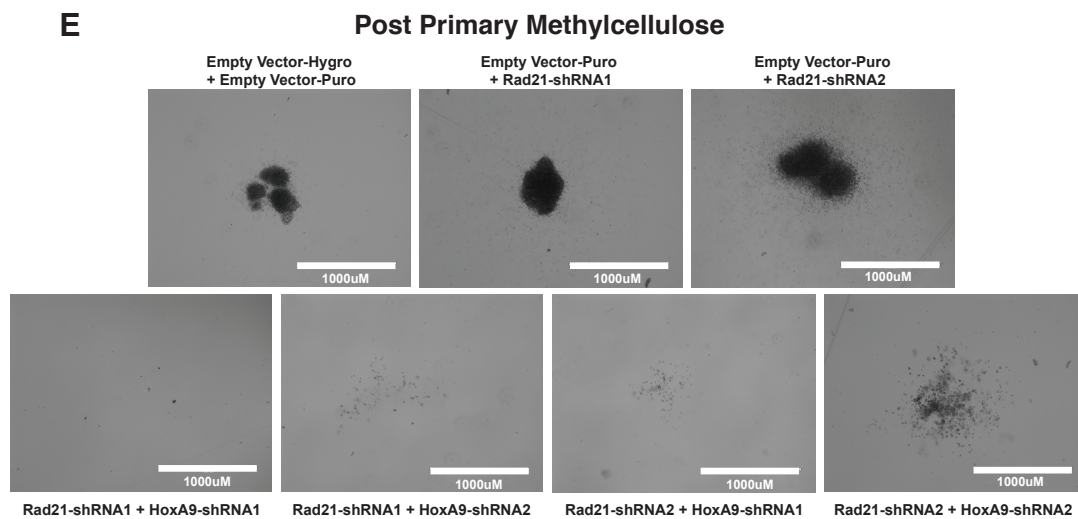
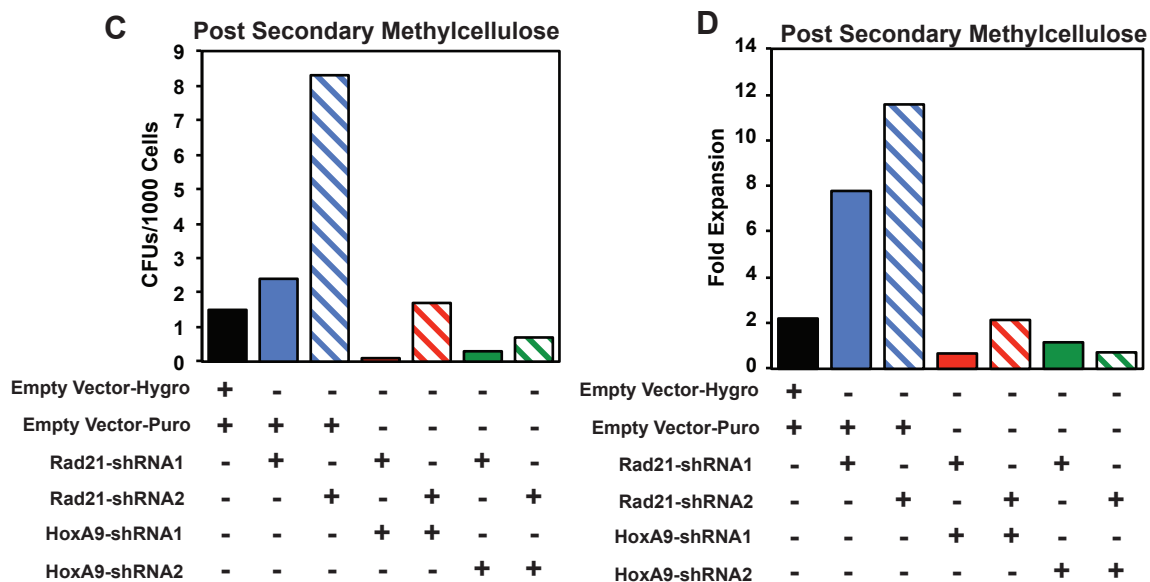
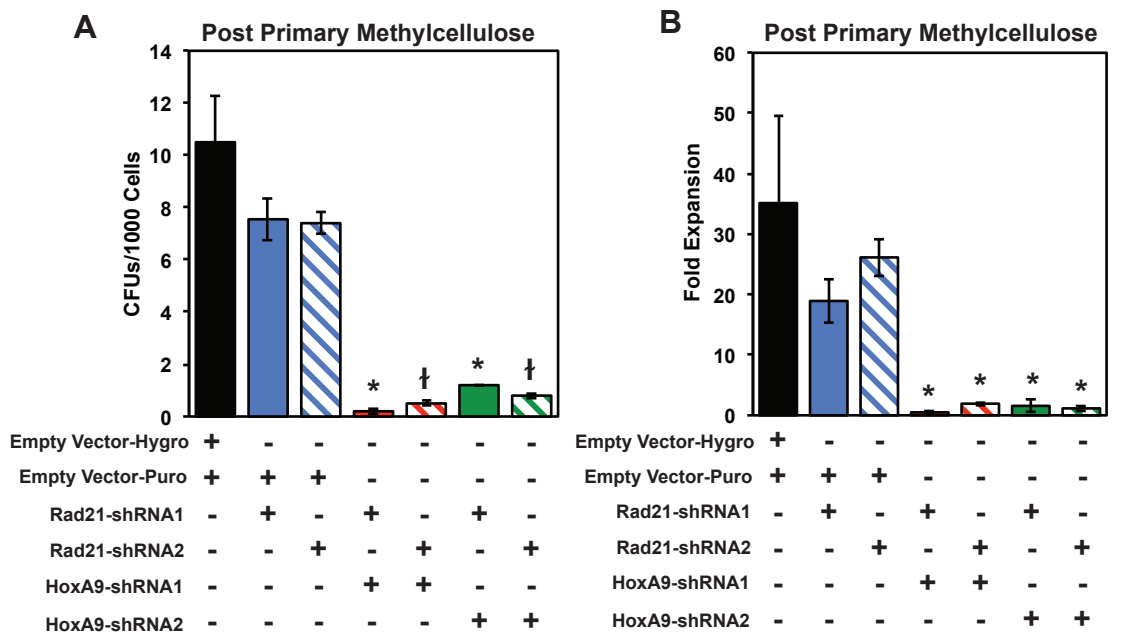


Figure S7. Co-depletion of *Hoxa9* and *Rad21* blocks HSPC self-renewal. (A-D) Empty vector, *Rad21*-shRNA, and *Rad21*-shRNA+*HoxA9*-shRNA HSPCs were serially passaged in methylcellulose and after each passage colony number (A: post primary and C: post secondary CFUs/1000 cells) and total cell numbers (B: post primary and D: post secondary fold expansion) were determined. (E) Representative images of empty vector, *Rad21*-shRNA, and *Rad21*-shRNA+*HoxA9*-shRNA treated HSPCs at the end of primary plating. Error bars represent standard error of the mean of 2 independent biological replicates. * = $p < 0.05$, † = $p < 0.01$ vs respective empty vector (IE *Rad21*-shRNA1+empty vector vs *Rad21*-shRNA1+*HoxA9*-shRNA1).

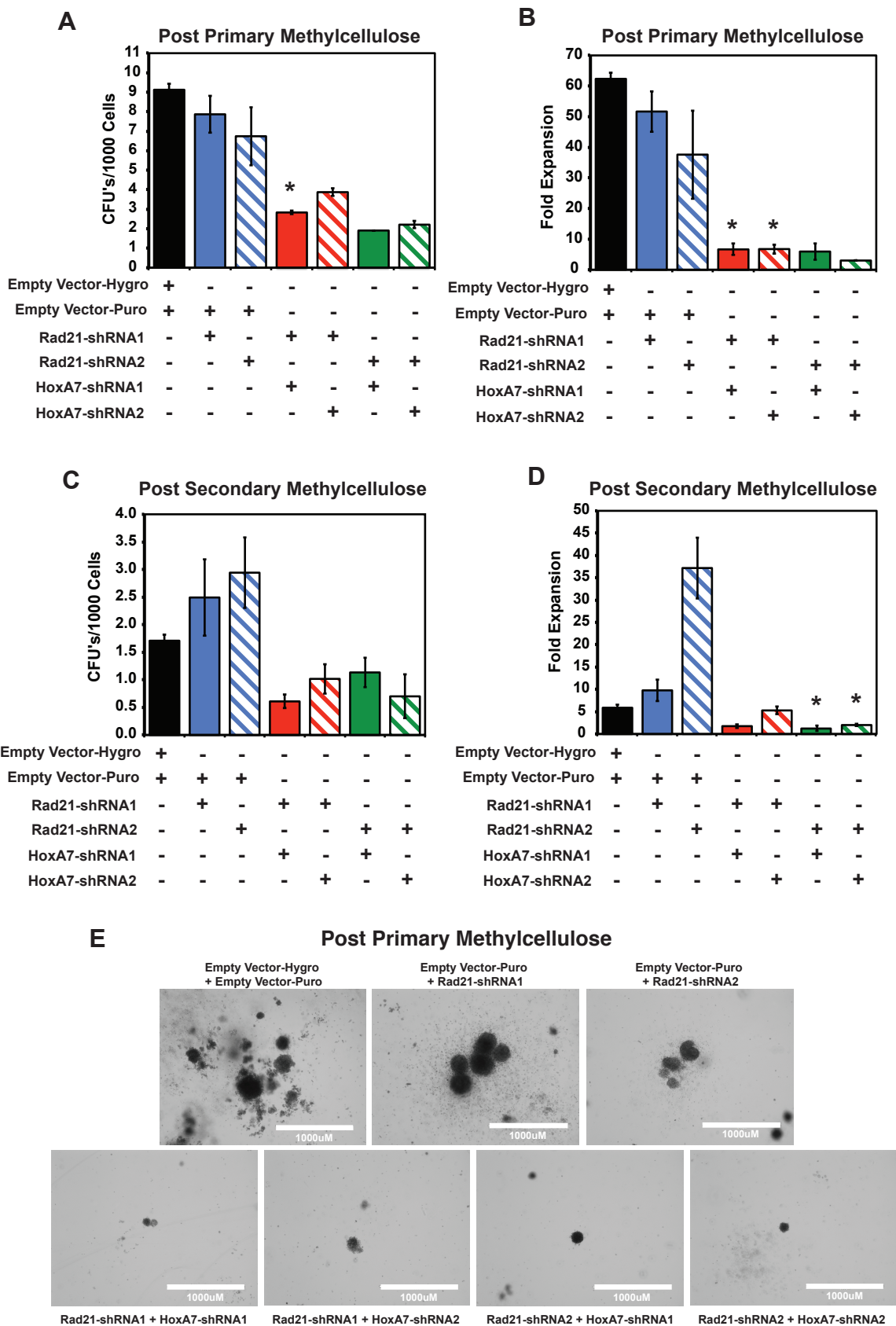


Figure S8. Co-depletion of *Hoxa7* and *Rad21* blocks HSPC self-renewal. (A-D) empty vector, *Rad21*-shRNA, and *Rad21*-shRNA+*Hoxa7*-shRNA HSPCs were serially passaged in methylcellulose and after each passage colony number (A: post primary and C: post secondary CFUs/1000 cells) and total cell numbers (B: post primary and D: post secondary fold expansion) were determined. (E) Representative images of empty vector, *Rad21*-shRNA, and *Rad21*-shRNA+*Hoxa7*-shRNA treated HSPCs at the end of primary plating. Error bars represent standard error of the mean of 2 independent biological replicates. * = $p < 0.05$ vs respective empty vector (IE *Rad21*-shRNA1+empty vector vs *Rad21*-shRNA1+*HoxA7*-shRNA1).

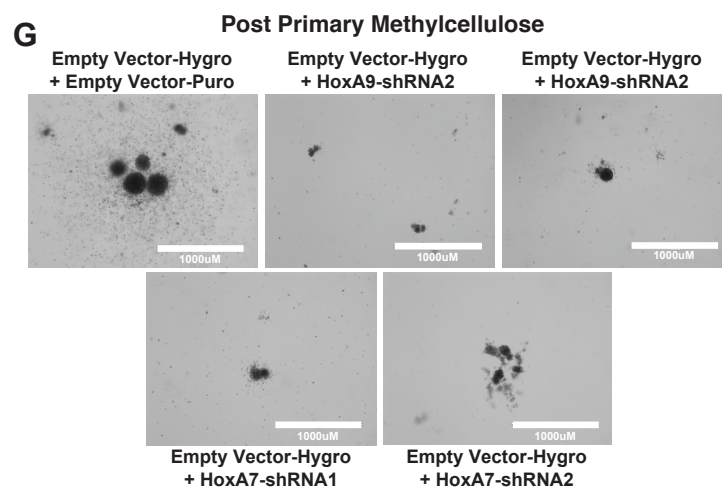
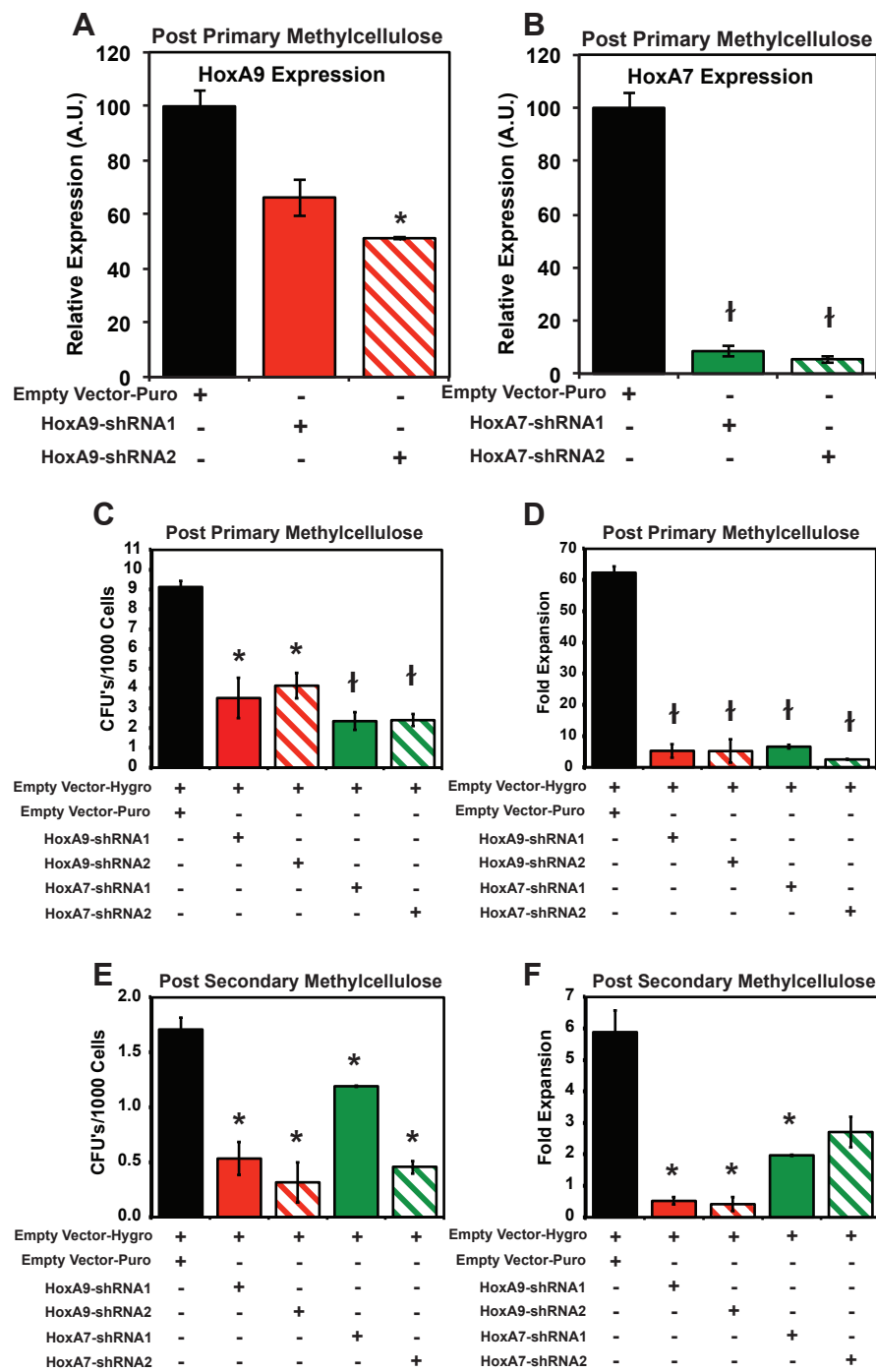
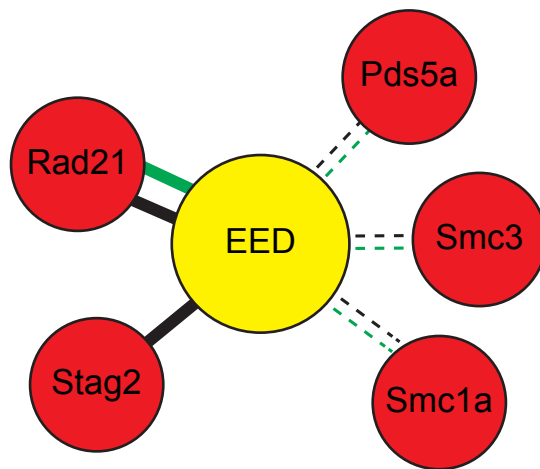


Figure S9. Depletion of *Hoxa7* and *Hoxa9* disrupts self-renewal in control HSPCs. (A) RT-qPCR analysis of *Hoxa7* or *Hoxa9*-depleted HSPCs at the end of primary methylcellulose culture showing shRNA-mediated depletion of *Hoxa7* and *Hoxa9*. (B-E) Empty vector, *Hoxa7*-shRNA, and *Hoxa9*-shRNA treated HSPCs were serially passaged in methylcellulose and after each passage colony number (B: post primary and D: post secondary CFUs/1000 cells) and total cell numbers (C: post primary and E: post secondary fold expansion) were determined. (F) Representative images of empty vector, *Hoxa7*-shRNA, and *Hoxa9*-shRNA treated HSPCs at the end of primary plating. The empty vector-puro+empty vector-hygro data points in panels A-D are identical to those shown in Figure S7A-D and were included here for comparison. Error bars represent standard error of the mean of 2 independent biological replicates.



— ESC (high)
— XEN (high)
- - - - - ESC (low)
- . - . - XEN (low)

Figure S10. Rad21 interacts with members of the PRC2 complex. Proteomic analysis indicating that the PRC2 subunit Eed directly interacts with members of the cohesin complex. Data were reanalyzed from {Cao:2014ki}. Interactions between Eed and cohesin complex members (including Rad21) were identified in both embryonic stem cells (ES) and extraembryonic endoderm (XEN) cells.

Mouse ENCODE Data

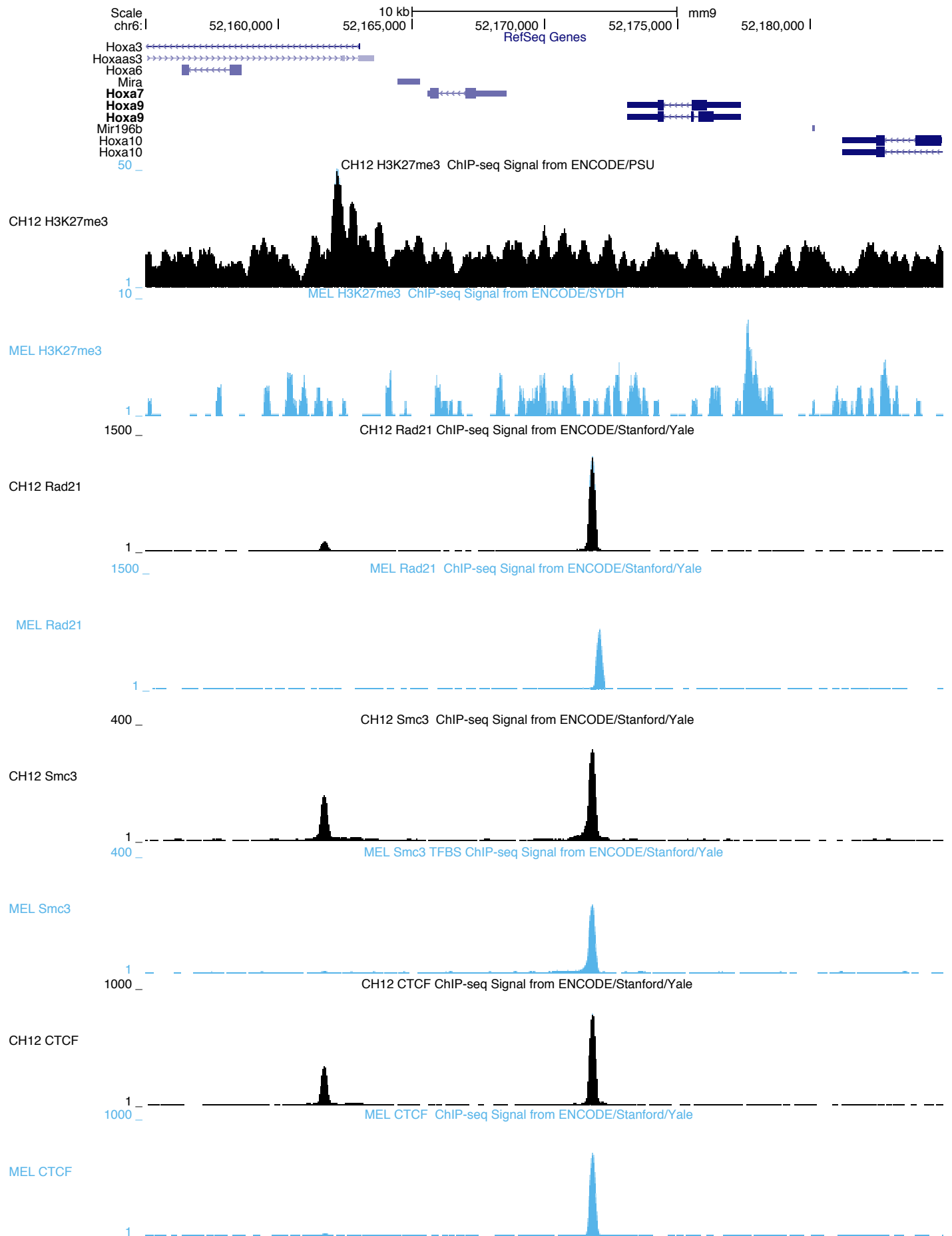


Figure S11. Cohesin occupies the *Hoxa9* locus in murine cells. Data from the ENCODE project was reanalyzed to monitor cohesin occupancy at the *Hoxa9* locus. Tracks show cohesin and Ctf peaks located within the murine *Hoxa9* locus in murine erythroleukemia cells (MEL) and lymphoma (CH12) cell lines. Tracks for histone H3K27me3, Rad21, Smc3, and Ctf are shown for each. For all tracks, the Y-axis scale is the same for each factor in the different cell lines, with the exception of H3K27me3.

Human ENCODE

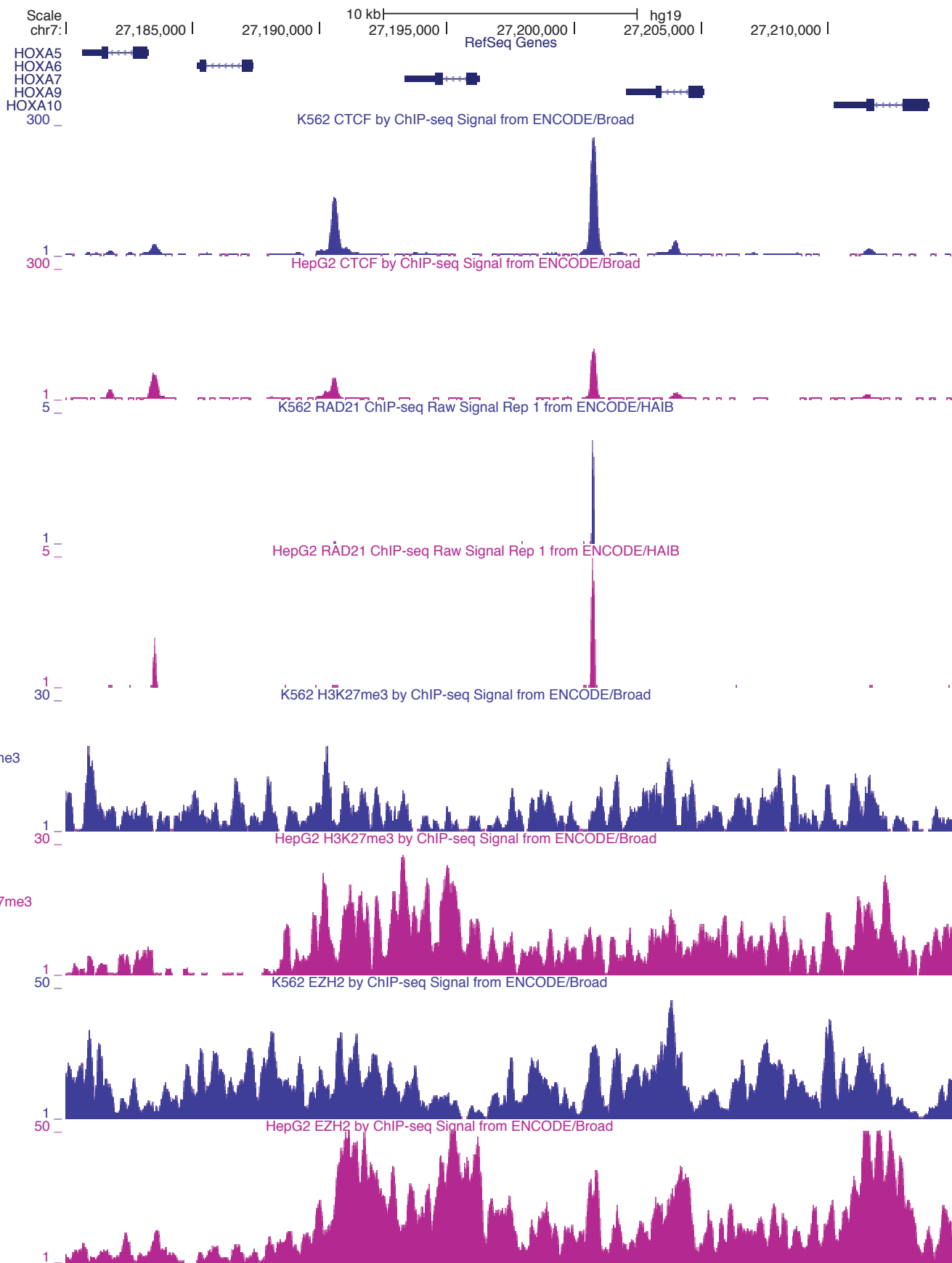


Figure S12. Cohesin occupies the *HOXA9* locus in human cells. Data from the ENCODE project was reanalyzed to monitor cohesin occupancy at the *HOXA9* locus. Tracks show cohesin and CTCF peaks located within the human *HOXA9* locus in myeloid leukemia cells (K562) and hepatocellular carcinoma (HepG2) cells. Tracks for histone H3K27me3, RAD21, EZH2, and CTCF are shown for each. For all tracks, the Y-axis scale is the same for each factor in the different cell lines.

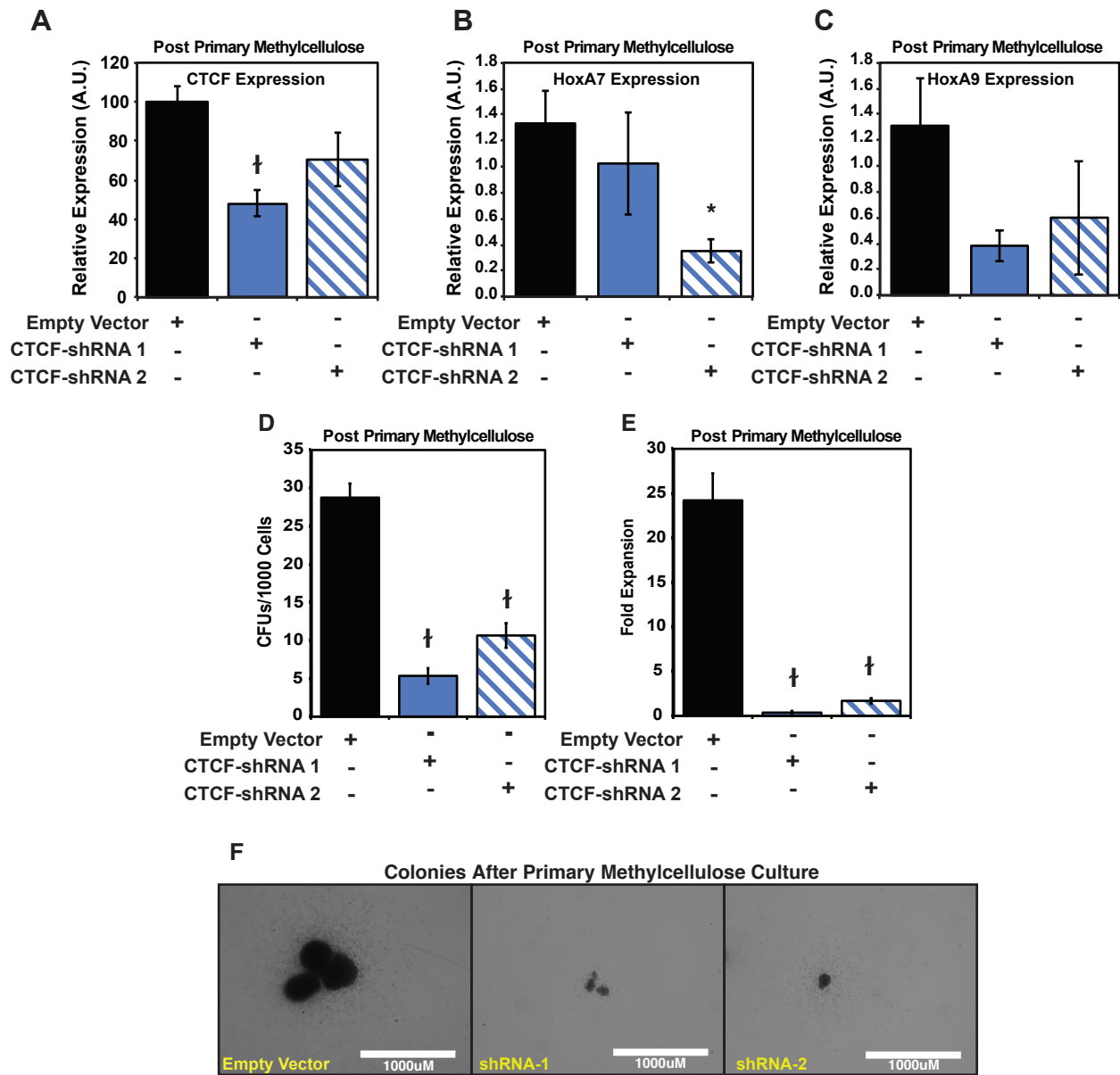


Figure S13. Depletion of *Ctcf* does not enhance HSPC self-renewal nor alter *Hoxa7* or *Hoxa9* expression. (A-C) RT-qPCR analysis of empty vector, or *Ctcf*-shRNA treated HSPCs the end of primary methylcellulose culture showing shRNA-mediated *Ctcf* depletion (A), and *Hoxa7* (B) and *Hoxa9* (C) expression. (D+E) Empty vector, and *Ctcf*-shRNA treated HSPCs were cultured in methylcellulose and at the end of primary passage, colony number (D; CFUs/1000 cells) and total cell numbers (E: fold expansion) were determined. (F) Representative images of empty vector, and *Ctcf*-shRNA treated HSPCs at the end of primary plating. Error bars represent standard error of the mean of 4 independent biological replicates.

Supplemental Tables:

Table S1: List of all qPCR primers used in these studies

Table S2: List of all antibodies used in these studies

Table S3: List of all publicly available datasets reanalyzed for this manuscript

Table S4: Transcript abundance in FPKM (Fragment per Kilobase per Million reads) for well-annotated RefSeq protein coding genes. Original data derived from{Mullenders:2015fi}.

Table S5: Microarray data from cells infected with an empty vector or two different shRNAs to *Rad21* at the end of primary passage.

Table S6: All gene sets enriched (p-value<0.01, FDR<25%) in *Rad21* depleted cells from GSEA.

Author Contributions:

The majority of experiments were performed by JF, JP, MR and CS. ZJG performed co-immunoprecipitation experiments, AMA assisted with flow cytometry approaches. KP and SR performed transcriptome analyses. JMS and CS analyzed TCGA data. MM, S. Malarkanna, JDC, and S. Milanovich assisted with the design and interpretation of the studies. All authors assisted with drafting and revising the manuscript, and approved with submission of the manuscript. S.Rao as corresponding author had full access to all raw data contained within this manuscript.

Acknowledgements:

S. Malarkannan is supported by NIH R01 AI102893, NCI R01 CA179363, HRHM Program of the MACC Fund, Nicholas Family Foundation; Gardetto Family Foundation; J. Crispino is supported by the Heartland Blood Centers and NIH R01 CA101774; M. McNulty is supported as a trainee of T32CA009560; S. Milanovich is supported by grants from National Institute of General Medical Sciences: NIH grant 5P20GM103548-5, and the MACC Fund.